DTU Library

Location-based restoration mechanism for multi-domain GMPLS networks

Manolova, Anna Vasileva; Calle, Eusibi; Ruepp, Sarah Renée; Marzo, Jose; Dittmann, Lars

Published in:
Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Location-based Restoration Mechanism for Multi-Domain GMPLS Networks

Anna Manolova, Sarah Ruepp, Lars Dittmann
E-mail: {anva, srru, ladd}@fotonik.dtu.dk
DTU Fotonik Technical University of Denmark
Ørsteds Pl. 343/130, 2800 Kgs. Lyngby, Denmark

Eusebi Calle, Jose Marzo
E-mail: {eusebi.calle, joseluis.marzo}@udg.edu
Institute of Informatics and Applications (IIiA)
University of Girona, 17071 Girona, Spain

***Abstract*—In this paper we propose and evaluate the efficiency of a location-based restoration mechanism in a dynamic multi-domain GMPLS network. We focus on inter-domain link failures and utilize the correlation between the actual position of a failed link along the path with the applied restoration technique. Our results show, that without violating the strong privacy preservation requirements between domains, the proposed mechanism has improved performance in terms of successfully restored connections compared to the traditional local-to-egress and end-to-end restoration approaches. Furthermore, the applied technique improves the availability of the connections by minimizing the restoration time.**

***Index Terms*—multi-domain network, GMPLS, restoration, location-based recovery.**

I. Introduction

NEW network technologies enable increasingly higher volumes of information. As networks grow, offering better quality of service, the consequences of a failure become more pronounced and thus network reliability has become a key requirement for next generation high speed networks. Network reliability can be provided through different fault management mechanisms applied at different network levels and time scales. A crucial aspect in developing a fault management system is the creation and routing of backup paths. This can be achieved either statically or dynamically. In the static case (protection), the connection backups are pre-established. In the dynamic case (restoration), the backups are created and/or routed dynamically in reaction to network faults. Protection schemes provide faster recovery of the failed path but require preplanned and/or pre-established protection paths, which increase the needed capacity in the network. Furthermore, protection schemes are applied only in static networks where the traffic demands are known. Restoration techniques on the other hand do not require additional installation of protection resources. They rely on the existing infrastructure and utilize the available spare capacity pool at the time of the failure.

With the increase of the available transmission capacity and the advances in providing new transport services, new requirements for failure handling have emerged. In particular the requirements for differentiated reliability provisioning have attracted attention and have resulted in extensive research work [1], [2]. Providing differentiated failure handling can be based not only on the service type but on other factors as well. The authors of [3] correlate the impact of a failure with the position of the failed link on the path and propose a novel routing mechanism which improves the availability of the established connections.

Unlike most of the research work in the area of differentiated reliability we focus on multi-domain link failure scenarios due to the unique implications the multi-domain environment pose on the reliability mechanisms. In our work we propose and investigate the efficiency of a position-based failure handling technique, which provides differentiation based on the actual position of the failure along the path of an affected connection. The work is focused on survivability in multi-domain single-layer connection-oriented network such as Generalized Multi-Protocol Label Switched (GMPLS) [4] networks. The main goal is to use a simple relation between the position of a failure and the applied failure handling so that the a differentiated failure recovery is performed. By this, an improvement of the availability of the Label Switched Paths (LSPs) is achieved without requiring additional capacity installation in the network. Furthermore, due to the confidentiality preservation policies between the domains in a multi-domain network, the used information for this process is as little as possible so that no network state or topology information is shared.

II. Survivability in multi-domain networks

The area of single domain survivability has been heavily investigated during the past years. The multi-domain case on the other hand has not received much attention. Only few literature sources focus on the problem of multi-domain survivability mainly because it is assumed that a failure in a given domain should be handled within the domain [5]. Furthermore, it is considered that the resilience mechanisms/principles should not be any different than in the single domain case.

A classification of the multi-domain resilience problems can be found in [5]. Three specific failure cases pose special challenges in multi-domain networks: inter-domain link failure, border node failure and a full domain failure. In these cases the survivability mechanism involves more than one domain and thus, requires coordination between the domains, additional intelligence and protocol extensions.

In the context of multi-domain restoration several issues can be outlined. The first problem comes from the limited visibility of the nodes regarding the multi-domain connectivity and the full network topology. The requirement for preservation of topological and state information within the domain is based on the strong confidentiality preservation policies between domains. This topology filtering has a very important implication - a border node in one domain would generally have information only about reachable destinations (edge nodes[1]) in the neighboring domain, but no information about core or border nodes. Thus, unless there are parallel links between border nodes there is no possibility to apply local restoration techniques for failed inter-domain links.

The next obstacle is the routing protocol used in the multi-domain environment. Currently there are no standards for multi-domain routing in GMPLS networks. Several approaches are being evaluated: the Border Gateway Protocol (BGP) [6], the Path Computation Element architecture (PCE) [7] and the E-NNI routing specification (only for intra-carrier application) [8]. Applying BGP necessitates protocol re-convergence after the failure in order to obtain the restoration path, since BGP provides only one path per destination. Employing a PCE for path computation also includes delay for restoration path computation, but the approach has the advantage of providing constrained-based path computation meeting various QoS requirements. The E-NNI approach also necessitates protocol re-convergence at different levels of the applied hierarchy. It is obvious that the slower the path computation is, the longer it takes to restore a failed connection. A possible solution is relying on pre-computed link_disjoint backup paths. Both the PCE approach and some BGP extensions [9] offer solutions for obtaining such paths.

The last obstacle can be found in the typical confidentiality preservation policies regarding the network state. Since we consider inter-domain link failures (the intra-domain failures are handled within each domain privately) it is still unclear how much information the domains are willing to share with their neighbors regarding the inter-domain links and how far the failure notification should be propagated. A typical approach is to confine the failure notification only within the domains attached to the failed link [5].

III. Improving connection availability

In order to evaluate how resilient a network is, one uses the notion of availability. A connection can be seen as a composite system where its components are the fiber optic cables, transceivers, amplifiers and other equipment employed for transporting traffic through the network. As such, its availability can be computed statistically based on reliability data, i.e. the failure frequency and failure repair rate of its components, measured over long periods of time [10]. Two such reliability measures are the Mean Time to Repair (MTTR), and the Mean Time Between Failures (MTBF). The MTTR is the average time spent performing corrective actions; during that period of time the affected component is non operative or "down". Considering connection oriented networks such as GMPLS, the availability of an established Label Switched Path (LSP) will depend on the applied recovery mechanism [11]. Three main recovery strategies exist. Fig. 1 a) depicts an end-to-end (E2E) recovery, also called path recovery. In this case, once the failure is detected a notification is sent to the node responsible of the recovery (in this case node 1, ingress node) and an E2E disjoint path is used to send the traffic. This scheme provides global path recovery, but during the recovery process there are traffic losses. This is due to the fact that the failure notification has to be passed from the node detecting it (node 5) to the ingress node (node 1). To avoid this notification causing data losses, local recovery, described in Fig. 1 b) is used. In this case the path is not protected end-to-end. Instead, each link along the path is protected separately. A combination of these two mechanisms is referred to as local-to-egress (L2E), Fig. 1 c). In multi-domain failure scenarios generally only the E2E and the L2E schemes can be applied due to the limited topological visibility of the border nodes. As explained in Sec. II local recovery is possible only if there are parallel links between border nodes. Furthermore, applying L2E recovery does not require traffic merging capabilities in each intermediate node.

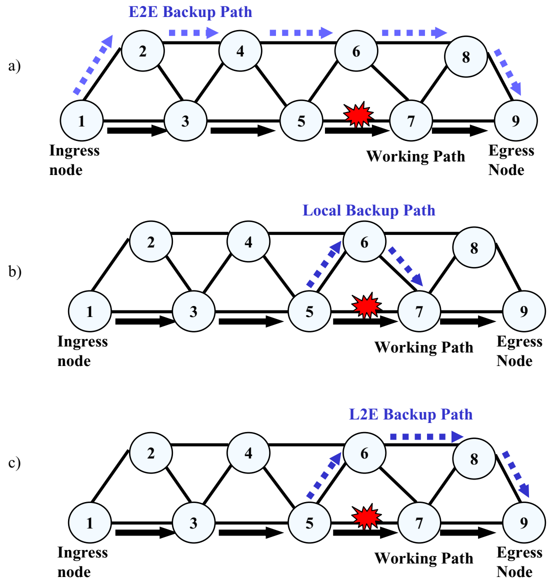


Fig. 1. Recovery schemes: a) End-to-End (path) recovery, b) Local recovery, c) Local-to-Egress recovery.

A. Availability analysis for different restoration techniques

The availability of a path can be expressed by either equation 1 or equation 2.

$$A_{path} = 1 - \frac{MTTR}{MTBF}, \qquad (1)$$

where MTTR is the Mean Time To Repair the failure and MTBF is the Mean Time Between Failures (we assume a constant value for all links in the network).

[1]In WDM transport networks the Edge nodes, communicating with the client networks, are considered to be the reachable destinations, disseminated by the inter-domain routing protocol.

$$A_{path} = \prod_i A_i = \prod_1 \{1 - \frac{MTTR_i}{MTBF_i}\}, \quad (2)$$

where A_i is the availability of each link along the path.

From the equations it can be seen that the availability of a LSP in case of restoration depends mainly on the MTTR value (for the overall path or for the separate links). The MTTR in case of restoration can be expressed by:

$$MTTR = T_d + T_n + T_{setup} + T_{sw}, \quad (3)$$

where T_d is the time to detect the failure, T_n is the time to notify the node, responsible for the failure recovery, T_{setup} is the time to set up the new LSP and T_{sw} is the time to switch over the traffic to the new established path.

Typically, the value of MTTR is dominated by the T_n and T_{setup}. It can be seen that depending on the applied restoration technique, the topology and connectivity of the network and the positioning of the failure along the path, the MTTR of a connection will be different. Thus, in this paper we propose to use differentiated treatment of affected LSPs based on the actual positioning of the failure along their path, in order to improve the availability of LSPs by minimizing the MTTR.

Fig. 2 illustrates the problem in a single domain environment in order to simplify the presentation of the concept. Depending on the position of the failure a certain restoration mechanism is more appropriate then others. In our analysis we assume that T_n and T_{setup} are proportional to the number of hops on the path and thus, we express them in terms of number of hops. For example, for failure 1 on Fig. 2 the connection between nodes 1 and 5 is more appropriate to be restored end-to-end, because $T_n^{e2e} + T_{setup}^{e2e} < T_n^{l2e} + T_{setup}^{l2e}$. On the other hand, for failure 2 the more appropriate restoration technique would be L2E.

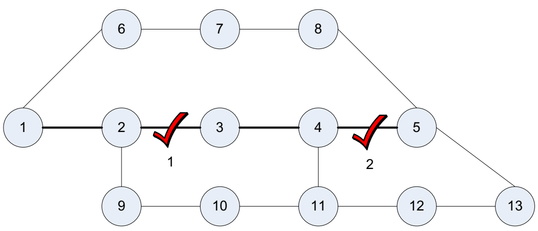


Fig. 2. Example of location dependent restoration.

IV. Location-based restoration for multi-domain GMPLS networks

In this section the general framework for a location-based restoration in connection-oriented multi-domain networks is outlined. An analysis of the challenges is given as well as proposals for their solution. Our work is explicitly focused on the case of inter-domain link failures.

A. Location-based LSP restoration

In location-dependent restoration the node upstream a link failure must take a decision: shall it restore a failed connection using the L2E technique or shall it signal the failure upstream to the head-end so that E2E restoration can be applied [2]. In order to take the most appropriate decision in our location-based restoration scheme a node must know its position in the overall path. We call this scheme Simple Location-Based Restoration (SLBR). Assuming equal times for failure notification and resource reservation per hop[3], the delay for recovery (i.e. the MTTR) can be expressed in terms of the distance $D(i, j)$ between the node detecting the failure i and the source/destination node j. Since the involved calculations are not complex (i.e. the packet processing per node is not a dominant factor) the delay for setting up a restoration connection is dominated by the propagation delay for the signaling packets. Thus, we express the distance between i and j in terms of number of hops. The SLBR technique requires the following decision to be made at the node detecting the failure:

- SLBR:
  $if \quad D(i,d) < D(i,s) \quad apply \quad L2E$
  $if \quad D(i,d) > D(i,s) \quad apply \quad E2E$
  $if \quad D(i,d) = D(i,s) \quad apply \quad random\ selection$
  where s denotes the source node, d the destination node and i the node upstream the link failure which detects it first.

If the node detecting the failure takes the decision to apply E2E, then the failure notification is propagated all the way back to the head-end of the connection where an E2E restoration is applied. The needed information for taking the decision in case of a failure can be provided using extended routing protocols and piggybacking the existing provisioning protocols. The following subsections discuss the challenges of the proposed mechanisms and possible solutions.

B. Challenges and solutions

As previously said, applying the proposed restoration mechanism necessitates the distribution of additional information and the application of specific path computation methods.

The first issue to be solved is related to the application of the multi-domain routing protocol. If standard BGP is used for multi-domain routing, the node detecting the failure must wait until the protocol re-converges to a stable solution in order to use a new path to restore the affected LSP. If a standard PCE architecture is used this requires either at least two parallel connections between domains [12], or an additional path computation delay if the Backward Recursive Path Computation technique is used dynamically [13]. In all described cases the restoration time (i.e. the MTTR) will be very long. In order to avoid this additional delay, and not to bound the

[2]See section II for explanation why local restoration has restricted application for inter-domain link failure in multi-domain networks.

[3]For increased accuracy more information can be added in the applied formulas.

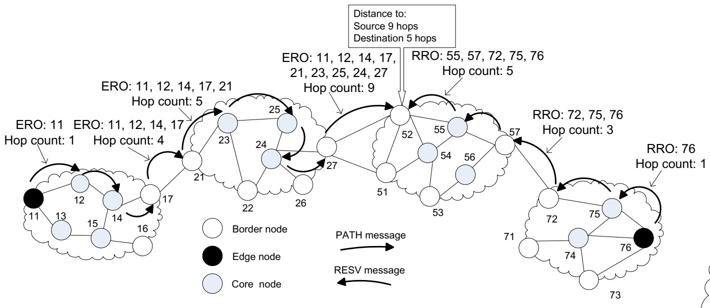


Fig. 3. Example of obtaining the exact position of a node along the LSP path.

network topology having double connections between domains a modified BGP protocol can be used, which provides more than one path per destination beforehand. Several solutions for multi-path dissemination in BGP networks exist [9], [14]. Another solution is the application of the E-NNI specification [8], but this requires OSPF-TE to be deployed in all domains. Whatever the chosen mechanism for multi-domain routing is, it is necessary that each border node has at least two disjoint paths per destination. In this way, the additional delay for restoration path computation will be avoided and the recovery time will mainly depend on the length of the restoration path and the distance from the failure to the point of recovery.

The second issue to be addressed is how to obtain the information, regarding the actual position of a border node along the overall path of the LSP. Several approaches can be taken. Fig. 3 illustrates two of them. The first option is using the Record Route Object (RRO) and the Explicit Route Object (ERO) during path setup. The ERO will typically carry the full path from the source up to the current border node. The RRO will carry the full path from the destination up to the current border node. In this way a border node will know exactly how far away it is from both the source and the destination. This information is stored in the border node per established LSP and is used in case a link attached to the node fails. A possible drawback is the confidentiality preservation requirements between ASes according to which no topological information should leak out the domain borders. The exact RRO and ERO objects carry sensitive topological information and are typically either erased when crossing borders or encoded [15]. In such a case a simple counter (TTL-type) can be used as depicted on Fig. 3, which keeps track of the number of nodes and can be added to the PATH and RESV messages of RSVP-TE [16]. An alternative approach is to use time-stamps but this necessitates time synchronization between the participating domains.

C. Operation example

In our implementation we use a modified BGP for obtaining two AS_disjoint paths per destination. Furthermore, we use a hop counter in the RSVP-TE messages for identifying the actual position of each border node along a path.

Fig. 4 illustrates the operation of the proposed SLBR scheme. Two options for the restoration of the affected LSP can be applied E2E or L2E (i.e. Local-to-Destination). It can

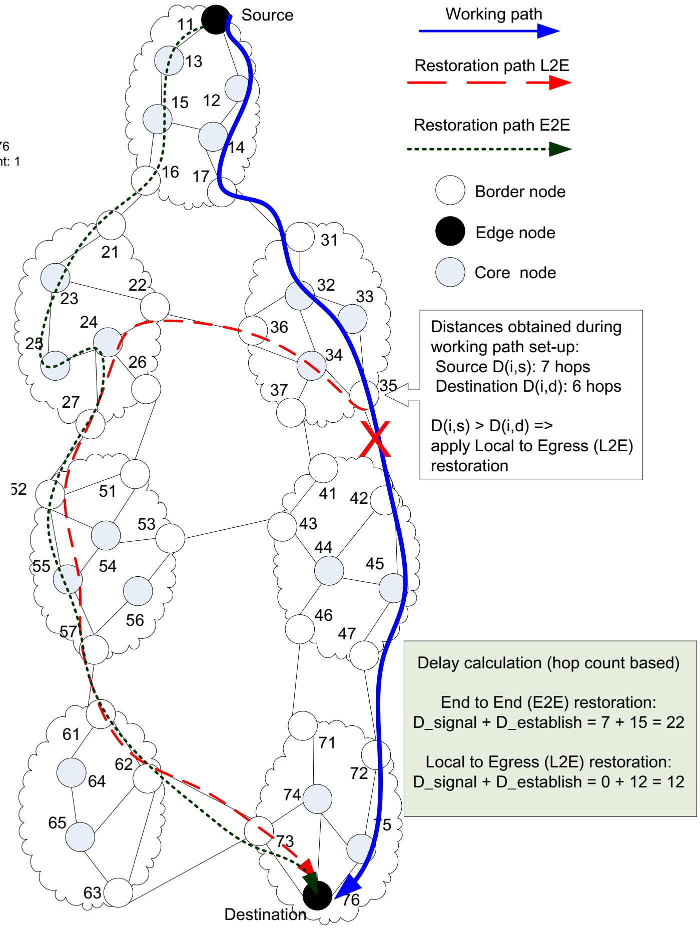


Fig. 4. Example of the operation of the SLBR mechanism.

be seen that if E2E restoration is applied, the setup delay will be proportional to 22 hops, and if L2E is applied - only 12 hops. Thus, if the proposed SLBR mechanism is applied, the affected LSPs which pass through the failed link will get differentiated handling and will be restored with a mechanism which seeks to minimize the MTTR.

V. Performance Evaluation

In order to evaluate the efficiency of the proposed SLBR scheme in a realistic dynamic multi-domain network we use the event-driven simulator tool OPNET [17]. A COST 266 Pan-European topology with 22 domains, 46 source/destination nodes and 42 bidirectional inter-domain links is used [18]. The internal topology of each domain is randomly generated. We make the following assumptions: each border and edge node has at least two AS_disjoint paths per destination[4]; all domains apply the proposed restoration scheme; the topology is flat, i.e. there are no client-server relationships between domains, each domain can be used as a transit domain; and in order to optimize the resource

[4] A modified BGP protocol is used for that purpose; all border and edge nodes are BGP speakers [9].

consumption within a domain the ingress border router is the point of repair for the cases when the L2E scheme is used.

Two are the main performance evaluation metrics we focus on: recovery success and recovery LSP setup delay (in our case equivalent to MTTR). The delay is computed as the time it takes for an LSP to be restored from the moment the link fails (i.e. it includes any propagation delays and packet processing delays).

A. Simulation setup

The simulated network has 50 wavelengths per link. During the provisioning phase no wavelength conversion is allowed. There are no parallel connections between border nodes, i.e. no local (link) restoration is possible. RSVP-TE is used as a reservation protocol and BGP as multi-domain routing protocol. Within each domain, shortest path routing is applied. During the recovery phase, new LSP requests continue to arrive in the system. The average connection duration is 600 sec., and the traffic is uniformly distributed. In case of a failure, span release is performed to free the unused resources from the affected LSPs.

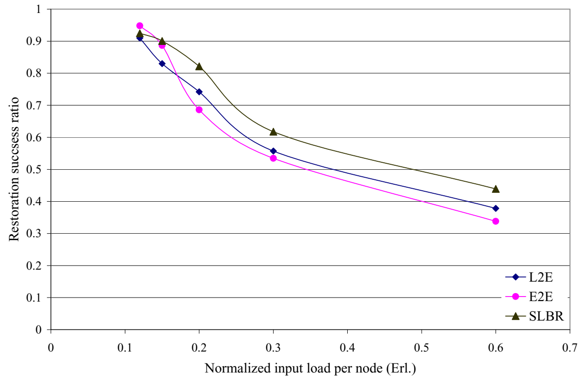
B. Results and future work

Two different setups are investigated in order to evaluate the efficiency of the proposed SLBR scheme. For the first setup, the restoration paths are allowed to seek for free wavelengths which might be different than the wavelengths of the working paths. This implies that at the point of local repair (if L2E restoration is performed) wavelength converters are required. Fig. 5 a) illustrates the average restoration success ratio for different traffic loads in the network. The input load per node is normalized to the capacity of the links. It can be seen that the proposed SLBR scheme outperforms the E2E and the L2E for medium and high loads. In the low load range the performance of the E2E scheme is better mainly due to the lack of contention for resources in the network and the fact that the E2E scheme offers better load balancing in the multi-domain environment.

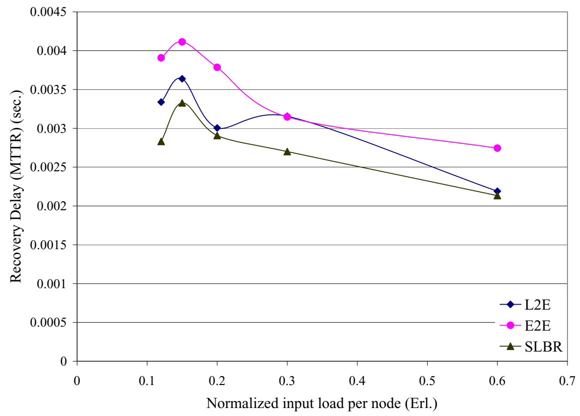
The results for the restoration path setup delays, depicted on Fig. 5 b), indicate that the higher the load in the network, the lower is the probability for a connection with long backup path to be successfully restored. Thus the setup delay for successfully restored connections is decreasing.

The next investigated performance metric is the amount of additional occupied resources for restoration in the network. For this, we compare the path lengths of the restored LSPs with their original length before failure. Fig. 6 a) shows the result. The E2E scheme yields the least amount of additionally occupied resources mainly due to the availability of equal length backup paths in the well-connected Pan-European topology. The proposed SLBR scheme has an intermediate position.

The irregularity of the depicted data for the low load range for both the setup delays and for the additional occupied resources is due to two reasons. First, the links in the network have different propagation delay. Second, at different loads and



(a) Recovery success ratio



(b) Setup delay for recovered LSPs

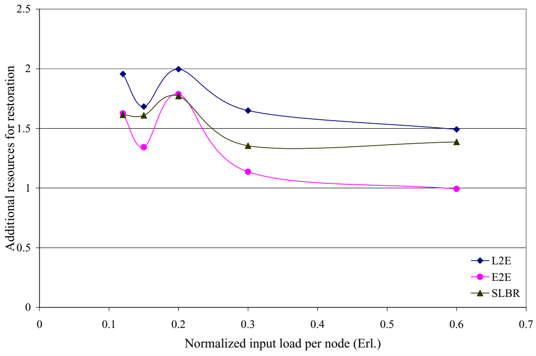
Fig. 5. Average recovery success ratio and setup delays for successfully recovered LSPs.

for different link failures different LSPs are affected which require backup paths with different length. Thus, a restoration path of 3 short hops may take less time to signal than a restoration path of 2 hops which have longer physical length.

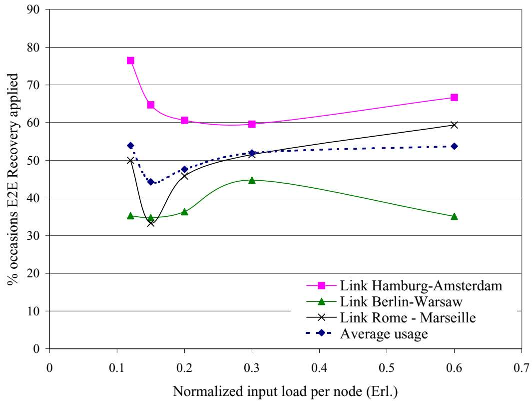
Since the SLBR scheme is a combination of the L2E and the E2E schemes it is interesting to see if the mechanism applies either of the schemes more than the other. Fig. 6 b) illustrates that depending on the particular link that fails, the percentage of cases where end-to-end restoration is applied under the SLBR scheme is different. This affects the performance of the SLBR scheme by binding it closer to the performance of the predominating restoration mechanism. On average though, about half of the affected connections are restored using E2E restoration under the SLBR scheme.

For the second setup, depicted on Fig. 7, the wavelength continuation constraint is applied, i.e. the restoration paths are required to take the same wavelengths as the working paths. In this case the SLBR scheme outperforms the other two only for the lowest load range. It is clearly more difficult and less probable to restore a failed connection when the availability of wavelengths is so severely reduced. These results show that in multi-domain networks the availability of limited wavelength conversion capabilities at border nodes for restoration purposes can be highly beneficial.

In our previous work [11] results show that there is a close

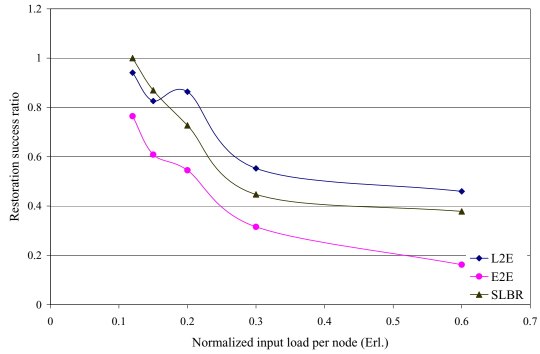


(a) Additionally occupied resources for recovery

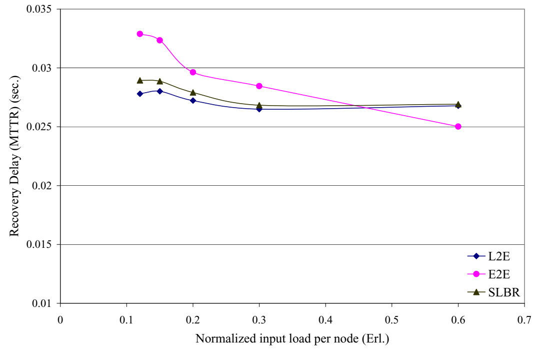


(b) Percentage of failed LSPs recovered end-to-end with SLBR scheme for three different link failures

Fig. 6. Additional resource occupation during recovery and % of E2E recovered LSPs with SLBR scheme.



(a) Recovery success ratio



(b) Setup delay for recovered LSPs

Fig. 7. Average recovery success ratio and setup delays when wavelength continuity in restoration is enforced.

relationship between network topology parameters in the inter-domain case for different recovery schemes. Our current work confirms this. In our future work it is interesting to extend the work presented in [11] to the restoration method suggested in this paper for multi-domain case. Similar extension can be done with using different routing schemes. Another interesting point for future work is the identification of the failure impacts depending on the failed link (zone) of the network. There are papers modeling this impact but only for a single-domain case [3], [19].

VI. Conclusion

In this paper we propose and investigate the performance of a location-based restoration mechanism for multi-domain connection-oriented dynamic networks. The proposed SLBR mechanism utilizes a simple relation between the position of a failed link along the path of the affected connection and the applied restoration technique. By this, an improvement in the restoration success ratio and the MTTR of the affected connections is achieved. The SLBR provides increased reliability and availability in multi-domain networks without raising privacy and confidentiality concerns. The additional information the mechanism relies on can be derived without sharing sensitive topological and/or state information between domains. Furthermore, we show that applying limited wavelength conversion solely in the border nodes of the domains can significantly improve the reliability of the network.

References

[1] Luo, H., Li, L., Yu, H.: "Routing Connections With Differentiated Reliability Requirements in WDM Mesh Networks," IEEE/ACM Transactions on Networking, Vol. 17, No. 1, February 2009.

[2] Jukan, A., Monitzer, A., R. van AS, H.: "QoS-Restorability in optical networks," In Proc. 24th European conference in Optical Communications (ECOC), Vol. 1, pp. 711-712, 1998.

[3] Calle. E., Marzo. J., Urra. A., Fabrega. L.: "Enhancing Fault Management Performance of Two-step QoS Routing Algorithms in GMPLS Networks," In proceedings of ICC 2004, 20-24 June, 2004, Paris (France).

[4] Mannie, E.: "Generalized Multi-Protocol Label Switching (GMPLS) Architecture," RFC 3945, October 2004.

[5] Larrabeiti, D., Romeral, R., Soto, I., Uruena, M., Cinkler, T., Szigeti, J., Tapolcai, J.: "Multi-domain issues of resilience," Transparent Optical Networks, 2005, Proceedings of 2005 7th International Conference Volume 1, 3-7 July 2005 Page(s): 375 - 380 Digital Object Identifier 10.1109/ICTON.2005.1505825.

[6] Rekhter, Y. and Hares, S.: "A Border Gateway Protocol 4 (BGP-4)," RFC 4271, January 2006.

[7] Farrel, A., Vasseur, J.-P. and Ash, J.: "A Path Computation Element (PCE)- Based Architecture," RFC 4655 (Informational), August 2006.

[8] The OIF Forum, *External Network-Network Interface (E-NNI) OSPF - based Routing - 1.0 (Intra-Carrier) Implementation Agreemt*, Jan. 2007.

[9] Manolova, A., Ruepp, S., Buron, J., Dittmann. L.: "On the Efficiency of BGP-TE Extensions for GMPLS Multi-Domain Routing," 13th ONDM Conference, February 2009.

[10] Verbrugge, S., Colle, D., Demeester, P., Huelsermann, R., and Jaeger, M.: "General availability model for multilayer transport networks," Proceedings 5th IEEE Int. Workshop on Design of Reliable Communication Networks (DRCN), Naples, Italy, 2005.

[11] Segovia, J., Calle, E., Vila, P., Marzo, J., and Tapolcai, J.: ”Topology-focused availability analysis of basic protection schemes in optical transport networks,” Journal of Optical Networking, Vol. 7, No. 4, April 2008.
[12] Staessens, D., Colle, D., Lievens, I., Pickavet, M., Demeester, P.: ”Enabling High Availability over Multiple Optical Networks,” IEEE Communications Magazine, June 2008.
[13] Vasseur, JP., Zhang, R.: ”A Backward Recursive PCE-based Computation (BRPC) Procedure To Compute Shortest Constrained Inter-domain Traffic Engineering Label Switched Paths,” IETF Draft, draft-ietf-pce-brpc-09.txt, Exp. Oct. 2008.
[14] Kushman, N., Kandula, S., Katabi, D., and Maggs. B.: ”R-BGP: Staying Connected in a Connected World,” 4th USENIX Symposium on Networked Systems Design & Implementation, April 2007.
[15] Farrel, A., Bradford, R., Vasseur, JP.: ”RSVP Extensions for Path Key Support,” Internet Draft, Expires: September 2009, draft-ietf-ccamp-path-key-ero-04.txt.
[16] Awduche, D., et. al: ”RSVP-TE: Extensions to RSVP for LSP Tunnels,” RFC 3209, December 2001.
[17] OPNET Modeler, http://www.opnet.com
[18] Inkret, R. et al.: ”Advanced Infrastructure for Photonik Networks,” Extended final report of COST Action 266, available at http://www.ufe.cz/dpt240/cost266/.
[19] Rozycki, P., Korniak, J., Jajszczyk, A.: ”Failure Detection and Notification in GMPLS Control Plane,” Workshop on GMPLS Performance: Control Plane Resilience, 2007